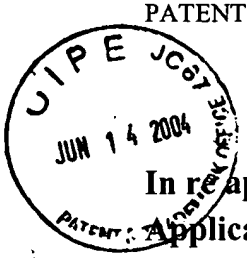
#16
6-21-04
P.2



IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re application of: Shum et al.
Application No. 09/338,176
Filed: June 22, 1999
Confirmation No. 1062
For: METHOD AND APPARATUS FOR RECOVERING A THREE-DIMENSIONAL SCENE FROM TWO-DIMENSIONAL IMAGES
Examiner: Allen C. Wong
Art Unit: 2613
Attorney Reference No. 3382-52053-01

CERTIFICATE OF MAILING

I hereby certify that this paper and the documents referred to as being attached or enclosed herewith are being deposited with the United States Postal Service as First Class Mail in an envelope addressed to: COMMISSIONER FOR PATENTS, P.O. BOX 1450, ALEXANDRIA, VA 22313-1450 on the date shown below.

Attorney for Applicant(s)
Stephen A. Wight, Attorney for Applicants

Date Mailed June 9, 2004

MAIL STOP APPEAL BRIEF – PATENTS
COMMISSIONER FOR PATENTS
P.O. BOX 1450
ALEXANDRIA, VA 22313-1450

REPLY BRIEF

Sir:

This Reply Brief is in response to the Examiner's Answer mailed April 21, 2004.

CONTENTS

I. REPLY TO EXAMINER'S ANSWER TO ARGUMENT

The pending claims 1-37 are allowable and should be passed to issue. The Examiner believes the rejections of the pending claims should be sustained. (*See,* Examiner's Answer "Answer" at p. 7.) However, Appellants are not persuaded by Examiner's answer. In fact, the Examiner has again failed to carry the burden of showing that the sole reference he relies upon, U.S. Pat. No. 6,046,745 to *Moriya et al.* (*"Moriya"*), teaches or suggests each and every element set forth in the pending claims.

A. The cited reference *Moriya* does not teach or suggest each and every element of independent claims 1, 23, 31, 36 and 37.

Claims 1-8

Claim 1 is directed to a method of generating a three-dimensional scene from a sequence of two-dimensional images. More particularly, claim 1 recites:

> A method of recovering a three-dimensional scene from two-dimensional images, the method comprising:
> providing a sequence of images;
> dividing the sequence of images into segments;
> performing three-dimensional reconstruction for each segment individually; and
> combining the three-dimensional reconstructed segments together to recover a three-dimensional scene for the sequence of images. (Emphasis added)

Sub-figures constituting a single basic figure as taught by Moriya do not teach or suggest a method of recovering a three-dimensional scene from a sequence of two-dimensional images by "providing a sequence of images" and "dividing the sequence of images into segments." The Examiner relies on element 4038 of FIG. 40 of *Moriya* described at col. 29, lns 28-31, which refers to "features of a basic figure 4038 constituting a three-dimensional model concerning sub-figures such as points, lines, planes, type of a cube, a classification of a linear line, and a curved line or the like." (*See, e.g.,* Answer, p. 5, lns. 14-20 alleging that *Moriya* at col. 29, lns 28-31 teaches or

suggests "dividing the sequence of images into segments"). The Applicants respectfully disagree.

Among other things, *Moriya* is directed to determining camera parameters and 3D-structure of an object by processing a single two-dimensional image of such an object. (*See, e.g.,* Appellants' brief pp. 7-8 for further proof of this assertion.) However, claim 1 recites "providing a sequence of images" and "dividing the sequence of images into segments." Appellants have made it clear that "a sequence of images", as clearly defined in the specification, is a series of multiple related images and the specification clearly teaches that it is this "sequence" that is divided into "segments", wherein "each segment contains a certain number of frames and the number of frames in each segment may vary in length." (*See,* specification p. 8, lns 8-10 defining the term segment; *See, also*, Appellants' Brief at pp. 8-9 for further support.) Thus, sub-figures "such as points, lines, planes, type of a cube, a classification of a linear line and a curved line or the like" constituting a single basic figure as taught by *Moriya* do not anticipate the claimed method comprising "providing a sequence of images" and "dividing the sequence of images into segments."

To further assert that *Moriya* anticipates "dividing the sequence of images into segments", in his Answer, the Examiner relies on *Moriya*'s col. 32, lns. 40-46, which states the following:

> In addition, in the present invention, since the camera parameters can be determined directly from a single image, it is not necessary to set or record the camera parameter when an image is actually taken with a camera, and further, it is also possible to apply CG modeling to any previously taken image (e.g., vintage or historical footage) of which camera parameters are not known. (Emphasis added)

In particular, the Examiner alleges that because *Moriya* uses the term "footage", it anticipates "dividing the sequence of images." As further support for his assertion, the Examiner relies on the following definition of the term "footage" in *Webster's II New Riverside University Dictionary*:

> *Footage-* A segment of motion-picture film, esp. a segment depicting a particular event or type of action.

Relying on this definition, the Examiner alleges "clearly, the discussion of footage discloses the sequence of images that are obtained from the segment of a motion picture film that depicts a particular event since it would take a multitude or sequence of images to capture the whole essence, scene of a particular event." (*See*, Answer at p. 6). However, for a reference to anticipate a claim, "the reference must describe the applicant's claimed invention sufficiently to have placed a person of ordinary skill in the field of the invention in possession of it." (*See*, *Schumer v. Laboratory Computer Systems, Inc.*, 308 F.3d 1304, 64 USPQ2d 1832 (Fed. Cir. 2002).) Also, "to anticipate, the reference must....enable one of ordinary skill in the art to make and use the claimed invention." (*See, Bristol Myers Squibb Co. v. Bridgewood Services, Inc.*, 246 F.3d 1368, at 1374, 58 USPQ2d 1508 (Fed. Cir. 2001).)

Here, *Moriya* refers to the term "footage" only once in the entire patent and uses that term unmistakably in the context of the term's meaning as "a single image" not "a sequence of images" as claimed. For instance, the passage in *Moriya* relied on by the Examiner says "it is also possible to apply CG modeling to any previously taken image (*e.g.*,...vintage footage)," which clearly shows that the term "footage" is being used as an exemplary source of a single image. Furthermore, *Moriya* offers no teaching or suggestion of how to process "a sequence of images" as claimed. Thus, a person of ordinary skill in art upon reading *Moriya* would not be led to "providing a sequence of images" and "dividing the sequence of images into segments" as claimed.

Furthermore, regardless of the meaning ascribed to the term "footage", the Examiner fails to show that *Moriya* teaches or suggests "dividing the sequence of images into segments", wherein "each segment contains a certain number of frames and the number of frames in each segment may vary in length." (*See,* specification p. 8, lns. 8-10 for clear definition of the term "segment.") As noted in the Appellants' Brief, a "segment" resulting from "dividing the sequence of images into segments" is clearly and

unambiguously defined in the specification as comprising "a certain number of *frames*" not "sub-figures such as points, lines, planes, type of a cube, a classification of a linear line and a curved line or the like" recited in *Moriya*. (*See*, Appellants' Brief at p. 9.) In fact, the above-mentioned definition of the term "footage" relied on by the Examiner itself suggests that an ordinary usage of the term "segment" refers to a segment comprising a "certain number of frames." For instance, by referring to "footage" as a "segment of a motion picture film" the Examiner's own statement seems to suggest that the ordinary meaning of the term "segment" in this context is "a certain number of frames" not the "sub-figures" of a single frame as taught by *Moriya*. (*See*, Answer at p. 6.)

Since *Moriya* fails to teach or suggest each and every element of claim 1, the rejection of claim 1 under 35 U.S.C. § 102(e) over *Moriya* is improper and claim 1 in its present form should be patentable. Claims 2-8 depend on claim 1 and at least for the reasons set forth above with respect to claim 1 they are also clearly patentable over the cited reference.

Claims 23-30

Amended claim 23 is also directed to a method of recovering a three-dimensional scene from a sequence of two-dimensional images. More particularly, claim 23 recites:

> A method of recovering a three-dimensional scene from a sequence of two-dimensional frames, comprising:
>
> (a) segmenting the sequence of two-dimensional frames;
>
> (b) identifying feature points in at least a first base frame in a first segment;
>
> (c) analyzing a second frame in the segment to identify the feature points in the second frame;
>
> (d) determining whether a threshold number of feature points from the base frame are identified in the second frame;
>
> (e) if a threshold number of feature points are identified in the second frame, adding the second frame to the segment; and
>
> (f) repeating (c) through (e) for subsequent frames until the number of feature points in a frame falls below the threshold number. (Emphasis added)

As amended, claim 23 recites "A method of recovering a three-dimensional scene from a sequence of two-dimensional frames, comprising: *segmenting the sequence of two-dimensional frames*." (Emphasis added) Thus, at least for the reasons set forth above with regard to claim 1, claim 23 and its dependent claims 24-30 should also be patentable.

Claims 31-35

Claim 31 is directed to an improvement in the method of recovering a three-dimensional scene from a sequence of two-dimensional frames. More particularly, claim 31 recites as follows:

> In a method of recovering a three-dimensional scene from a sequence of two-dimensional frames, an improvement comprising dividing a long sequence of frames into segments and reducing the number of frames in each segment by representing the segments using between two and five representative frames per segment, wherein the representative frames are used to recover the three-dimensional scene and remaining frames are discarded so that the three-dimensional scene is effectively compressed. (Emphasis added)

Claim 31 recites "In a method of recovering a three dimensional scene from a sequence of two-dimensional frames, an improvement comprising *dividing a long sequence of frames into segments.*" Thus, at least for the reasons set forth above with regard to claim 1, claim 31 and its dependent claims 32-35 should also be patentable.

Claim 36

Claim 36 is directed to a computer-readable medium having computer-executable instructions for performing a method of recovering a three-dimensional scene from a sequence of two-dimensional frames. More particularly, claim 36 recites as follows:

> A computer-readable medium having computer-executable instructions for performing a method comprising:
> providing a sequence of two-dimensional frames;

dividing the sequence into segments;
calculating a partial model for each segment that includes three-dimensional coordinates and camera pose for features within the frames;
extracting virtual key frames from each partial model, the virtual key frames having three-dimensional coordinates for the frames and an uncertainty associated with the frames; and
bundle adjusting the virtual key frames to obtain a complete three-dimensional reconstruction of the two-dimensional frames. (Emphasis added)

Claim 36 recites "providing a sequence of two-dimensional frames" and "dividing the sequence into segments." Thus, at least for the reasons set forth above with regard to claim 1, claim 36 should also be patentable.

Claim 37

Claim 37 is directed to an apparatus for recovering a three dimensional scene from a sequence of two-dimensional frames by segmenting the sequence of frames. More particularly, claim 37 recites as follows:

An apparatus for recovering a three-dimensional scene from a sequence of two-dimensional frames by segmenting the frames, comprising:
means for capturing two-dimensional images;
means for dividing the sequence into segments;
means for calculating a partial model for each segment that includes three-dimensional coordinates and camera pose for features within the frames;
means for extracting virtual key frames from each partial model; and
means for bundle adjusting the virtual key frames to obtain a complete three-dimensional reconstruction of the two-dimensional frames. (Emphasis added)

Claim 36 recites "means for dividing the sequence into segments." Thus, at least for the reasons set forth above regarding claim 1, claim 37 should also be patentable.

B. The cited reference, *Moriya* does not teach or suggest each and every element of independent claim 9.

Claim 9:

Claim 9 recites a method of recovering a three-dimensional scene from a sequence of two-dimensional images. More particularly, claim 9 recites as follows:

> A method of recovering a three-dimensional scene from two-dimensional images, the method comprising:
> identifying a sequence of two-dimensional frames that include two-dimensional images;
> dividing the sequence of frames into segments, wherein a segment includes a plurality of frames;
> for each segment, encoding the frames in the segment into at least two virtual frames that include a three-dimensional structure for the segment and an uncertainty associated with the segment.

Sub-figures constituting a single basic figure as taught by Moriya do not teach or suggest a method of recovering a three-dimensional scene from a sequence of two-dimensional images by "dividing the sequence of frames into segments, wherein a segment includes a plurality of frames." (Emphasis added) First of all, as noted above with reference to claim 1, the specification makes it amply clear that the claimed "segments" comprise "a certain number of *frames*." (*See,* specification p. 8, lns. 8-10.) However, the additional limitation of claim 9 which recites "wherein a segment includes a plurality of frames" makes it abundantly clear that the claimed "segments" comprise frames not "sub-figures such as points, lines, planes..." taught by *Moriya*. Thus, "dividing the sequence of frames into *segments, wherein a segment includes a plurality of frames"* cannot be anticipated by *Moriya*'s "sub-figures such as points, lines, planes, type of a cube, a classification of a linear line and a curved line or the like."

In his answer, the Examiner has chosen to simply ignore this additional limitation of claim 9, as if the Appellants' grouping of claims was superficial. In fact, the Examiner's answer to Appellants' arguments is identical to that of claim 1, without any acknowledgment of the differences between these two claims. At the risk of belaboring the point, "sub-figures such as points, lines, planes, type of a cube, a classification of a

linear line and a curved line or the like" are not "*segments, wherein a segment includes a plurality of frames.*" (Emphasis added)

Furthermore, even with his reliance on the term "footage" the Examiner fails to carry his burden of showing that *Moriya* anticipates "dividing the sequence of frames into segments, wherein a segment includes a plurality of frames." Even assuming for argument's sake what the Examiner alleges is true, at best, the term "footage" may suggest to one of ordinary skill in the art that a "sequence of frames" may be divided into "sub-figures such as points, lines, planes..." as taught by *Moriya*. However, that does not suggest "dividing the sequence of frames into *segments, wherein a segment includes a plurality of frames*." (Emphasis added)

A plain reading of claim 9 will not allow the interpretation relied on by the Examiner. Thus, the rejection of claim 9 under 35 U.S.C. § 102(e) over *Moriya* is improper for this reason in addition to the arguments set forth above with respect to claim 1.

II. CONCLUSION

In light of the arguments presented above the rejection of claims 1-37 should be reversed and all claims passed to issue.

Respectfully submitted,

KLARQUIST SPARKMAN, LLP

By _____
Stephen A. Wight
Registration No. 37,759

One World Trade Center, Suite 1600
121 S.W. Salmon Street
Portland, Oregon 97204
Telephone: (503) 226-7391
Facsimile: (503) 228-9446

APPENDIX A

CLAIMS ON APPEAL

1. A method of recovering a three-dimensional scene from two-dimensional images, the method comprising:

providing a sequence of images;

dividing the sequence of images into segments;

performing three-dimensional reconstruction for each segment individually; and

combining the three-dimensional reconstructed segments together to recover a three-dimensional scene for the sequence of images.

2. The method of claim 1 wherein performing includes creating virtual key frames for each of the segments, wherein the virtual key frames are only a subset of the images in a segment but are a representation of all of the images in that segment.

3. The method of claim 1 wherein the images contain feature points and a number of images included in a segment is based upon a number of feature points tracked between the images in that segment.

4. The method of claim 1 wherein performing further includes:

performing a two-frame structure-from-motion algorithm to create a plurality of local models for each segment; and

combining the plurality of local models by eliminating scale ambiguity.

5. The method of claim 4 further comprising:

bundle adjusting the combined local models to obtain a partial three-dimensional model for each segment;

extracting virtual key frames from the partial three-dimensional model, wherein the virtual key frames include three-dimensional coordinates for the images and an associated uncertainty; and

bundle adjusting all segments to obtain a complete three-dimensional model.

6. The method of claim 1 further including:

identifying feature points in the images;

estimating three-dimensional coordinates of the feature points; and

estimating a camera rotation and translation for a camera that captured the sequence of images.

7. The method of claim 1 wherein combining includes performing a non-linear minimization process across the different segments through bundle adjustment.

8. A computer-readable medium having computer-executable instructions for performing the method recited in claim 1.

9. A method of recovering a three-dimensional scene from two-dimensional images, the method comprising:

identifying a sequence of two-dimensional frames that include two-dimensional images;

dividing the sequence of frames into segments, wherein a segment includes a plurality of frames;

for each segment, encoding the frames in the segment into at least two virtual frames that include a three-dimensional structure for the segment and an uncertainty associated with the segment.

10. The method of claim 9 wherein dividing includes:

17. The method of claim 9 wherein encoding includes:

choosing at least two frames in the segment that are at least a threshold number of frames apart;

for each of the at least two chosen frames, projecting a plurality of three-dimensional points into a corresponding virtual frame; and

for each of the at least two chosen frames, projecting an uncertainty into the corresponding virtual frame.

18. The method of claim 9 further including bundle adjusting the virtual frames from the segments to create a three-dimensional reconstruction.

19. The method of claim 9 further including identifying feature points in the frames by using motion estimation.

20. The method of claim 19 wherein the motion estimation includes:

creating a template block in a first frame including a feature point and neighboring pixels adjacent the feature point;

creating a search window used in a second frame; and

comparing an intensity difference between the search window and the template block to locate the feature point in the second frame.

21. The method of claim 9 wherein at most two virtual frames are used.

22. A computer-readable medium having computer-executable instructions for performing the method recited in claim 9.

identifying a base frame;

identifying feature points in the base frame; and

defining the segments such that every frame in a segment has at least a predetermined percentage of feature points identified in the base frame.

11. The method of claim 9 wherein the segments vary in length and wherein the length is associated with the number of frames in the segment.

12. The method of claim 9 further including:

identifying feature points in the sequence of two-dimensional frames;

estimating three-dimensional coordinates for the feature points; and

estimating camera rotation and translation for the feature points.

13. The method of claim 12 wherein estimating the three-dimensional coordinates includes applying a two-frame structure-from-motion algorithm to the sequence of two-dimensional frames.

14. The method of claim 9 further including:

dividing a segment into multiple frame pairs;

applying a two-frame structure-from-motion algorithm to the multiple frame pairs to create a plurality of local models; and

scaling the local models so that they are on a similar coordinate system.

15. The method of claim 14 wherein each of the multiple frame pairs includes a common base frame and one other frame in the segment.

16. The method of claim 15 further including interpolating frames between the multiple frame pairs.

23. A method of recovering a three-dimensional scene from a sequence of two-dimensional frames, comprising:

(a) segmenting the sequence of two-dimensional frames;

(b) identifying feature points in at least a first base frame in a first segment;

(c) analyzing a second frame in the segment to identify the feature points in the second frame;

(d) determining whether a threshold number of feature points from the base frame are identified in the second frame;

(e) if a threshold number of feature points are identified in the second frame, adding the second frame to the segment; and

(f) repeating (c) through (e) for subsequent frames until the number of feature points in a frame falls below the threshold number.

24. The method of claim 23 further including designating a frame that falls below the threshold number as a base frame in a second segment and repeating (b) through (e) for the second segment.

25. The method of claim 23 further including performing motion estimation to identify the feature points.

26. The method of claim 23 further including using corners as the feature points.

27. The method of claim 23 wherein the number of frames varies between segments.

28. The method of claim 23 further including creating two virtual key frames per segment.

29. The method of claim 28 further including bundle adjusting the virtual key frames of all the segments to obtain a three-dimensional reconstruction.

30. A computer-readable medium having computer-executable instructions for performing the method recited in claim 23.

31. In a method of recovering a three-dimensional scene from a sequence of two-dimensional frames, an improvement comprising dividing a long sequence of frames into segments and reducing the number of frames in each segment by representing the segments using between two and five representative frames per segment, wherein the representative frames are used to recover the three-dimensional scene and remaining frames are discarded so that the three-dimensional scene is effectively compressed.

32. The method of claim 31 wherein each of the representative frames have an uncertainty associated therewith.

33. The method of claim 31 wherein the long sequence includes over 75 frames.

34. The method of claim 31 wherein dividing the long sequence into segments includes identifying a base frame and tracking feature points between frames in the

sequence and the base frame and ending a segment whenever a frame does not contain a predetermined threshold of feature points that are contained in the base frame.

35. The method of claim 31 further including performing a two-frame structure-from-motion algorithm on each of the segments to create a partial model.

36. A computer-readable medium having computer-executable instructions for performing a method comprising:

providing a sequence of two-dimensional frames;

dividing the sequence into segments;

calculating a partial model for each segment that includes three-dimensional coordinates and camera pose for features within the frames;

extracting virtual key frames from each partial model, the virtual key frames having three-dimensional coordinates for the frames and an uncertainty associated with the frames; and

bundle adjusting the virtual key frames to obtain a complete three-dimensional reconstruction of the two-dimensional frames.

37. An apparatus for recovering a three-dimensional scene from a sequence of two-dimensional frames by segmenting the frames, comprising:

means for capturing two-dimensional images;

means for dividing the sequence into segments;

means for calculating a partial model for each segment that includes three-dimensional coordinates and camera pose for features within the frames;

means for extracting virtual key frames from each partial model; and

means for bundle adjusting the virtual key frames to obtain a complete three-dimensional reconstruction of the two-dimensional frames.